Title	On Malign Input Distributions for Algorithms(Mathematical Logic and Applications'92)
Author(s)	Kobayashi, Kojiro
Citation	数理解析研究所講究録 (1993), 818: 20-34
Issue Date	1993-01
URL	http://hdl.handle.net/2433/83139
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第818巻 1993年 20-34

On Malign Input Distributions for Algorithms

東工大・理 小林孝次郎 (Kojiro Kobayashi)

1 Introduction

One important problem in the theory of complexity of computation is how the average-case computation time of algorithms is related to the probability distribution of inputs.

Concerning this problem, Li and Vitányi [2] found an interesting result that if inputs to algorithms are selected with the probability that is proportional to a semicomputable measure that is known as an a priori measure [3], then the worst-case computation time and the average-case computation time are roughly the same for any algorithm. A more precise statement of this result is as follows.

By a measure we mean a function μ from $\{0,1\}^*$ (the set of all binary sequences) to real numbers such that $\mu(x) \geq 0$ for any x and $\mu(\{0,1\}^*) < \infty$. For an algorithm A and a measure μ, let $t_A^{\mathrm{wo}}(n)$ denote the worst-case computation time of A for inputs from $\{0,1\}^n$ (the set of all binary sequences of length n) and let $t_A^{\mathrm{av},\mu}(n)$ denote the average-case computation time of A for inputs from $\{0,1\}^n$ where a sequence x in $\{0,1\}^n$ is given to A with the probability $\mu(x)/\mu(\{0,1\}^n)$. Li and Vitányi's result says that if μ is an a priori measure then for any algorithm A there exists a constant $c(>0)$ such that $t_A^{\mathrm{wo}}(n) \leq ct_A^{\mathrm{av},\mu}(n)$ for any n. Miltersen [4] called measures having this property "malign measures." Then Li and Vitányi's result says that a priori measures are malign.

A priori measures are very complicated mathematical objects and it is difficult to imagine a real situation where inputs are generated with the probability that is proportional to an a priori measure. Hence, if "a priori"-ness and malignness are equivalent, we may regard Li and Vitányi's result as one interesting but pathological phenomenon. In the present pa-

per, we consider the problem of whether the two notions "a priori"-ness and malignness are different or not.

For this problem, we first introduce the notion of strongly malign measures and point out that what Li and Vitányi proved is that a priori measures are strongly malign. Although the intuitive notion of malignness allows several natural variations of definition of malignness other than the one introduced by Miltersen, strong malignness seems to imply all of these variations. Hence we may regard strong malignness as the strongest of these variations. Then we show that there exists a semicomputable strongly malign measure that is not a priori. This means that the two notions "a priori"-ness and malignness are different in one strong sense.

We also show two results concerning a priori measures and Kolmogorov complexity. For a rational number p such that $0 < p < 1$, let $P_p(x)$ denote the probability that a universal prefix-free algorithm A_U outputs a binary sequence x when the input binary sequence to A_U is randomly selected with p as the probability that the symbol 1 is selected. It is known that $P(x) = P_{1/2}(x)$ is an a priori measure. We show that $P_p(x)$ is an a priori measure for each p. Let $H_p(x)$ denote the smallest of $|u|_p$ for binary sequences u such that A_U outputs x when u is given to A_U as an input, where $|u|_p$ denotes the value $(-\log_2(1-p))$(number of 0's in u) + $(-\log_2 p)$(number of 1's in u). The value $H(x) = H_{1/2}(x)$ is known as the (prefix-free algorithm based) Kolmogorov complexity of x. We show that for each p there exists a constant c such that $|H(x) - H_p(x)| \leq c$ for any x.

2 Preliminaries

Let Σ denote the set $\{0, 1\}$ of two symbols 0, 1, Σ^n denote the set of all sequences of 0, 1 of length n, and Σ^*denote the set $\Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \ldots$. We call an element of Σ^* a *word*. For each word x, let $|x|$ denote its length. Let λ denote the empty word, that is, the sequence of length 0. We use N, Q, R to denote the sets of all natural numbers, all rational numbers,

and all real numbers respectively. We assume that some standard one-to-one mapping from N onto Σ^* is fixed and we identify a number with its corresponding word. Throughout in this paper, $\log x$ means $\log_2 x$.

By a *measure* we mean a function μ from Σ^* to R such that $\mu(x) \geq 0$ for any x and $\mu(\Sigma^*) < \infty$ (for $L \subseteq \Sigma^*$, $\mu(L)$ denotes $\Sigma_{x \in L} \mu(x)$). A measure μ is said to be *semicomputable* if there exists a computable function f from $\Sigma^* \times$ N to Q such that $f(x,n) \leq f(x,n')$ for $n \leq n'$ and $\lim_{n \to \infty} f(x,n) = \mu(x)$. For two functions f, g from N to R, we write $f \lesssim g$ if there exists a constant $c(> 0)$ such that $f(n) \leq cg(n)$ for any n. We write $f \approx g$ if both of $f \lesssim g$, $g \lesssim f$ hold true.

By an *algorithm* we mean a deterministic Turing machine. We assume that algorithms accept words as inputs and generate words as outputs if they halt. Let $\phi_A(x)$ denote the output of an algorithm A for an input x. When the algorithm does not halt, $\phi_A(x)$ is undefined. We assume that some standard one-to-one mapping from the set of all (representations of) algorithms onto Σ^* is fixed and we identify an algorithm with its corresponding word.

By a *step counting function*, we mean any partial function $\mathrm{time}_A(x)$ that assigns a natural number to an algorithm A and an input x that satisfies the following two conditions (the "Blum's axioms" [1]): (1) $\phi_A(x)$ is defined $\Longleftrightarrow$ $\mathrm{time}_A(x)$ is defined; (2) the predicate "$\mathrm{time}_A(x) \leq t$" on A, x, t is decidable. We understand that the predicate "$\mathrm{time}_A(x) \leq t$" is false when $\mathrm{time}_A(x)$ is undefined.

Suppose that one step counting function is fixed. We define the worst case computation time $t_A^{\mathrm{wo}}(n)$ of an algorithm A for inputs of length n by $t_A^{\mathrm{wo}}(n) = \max_{x \in \Sigma^n} \mathrm{time}_A(x)$. For each measure μ, we define the average case computation time $t_A^{\mathrm{av},\mu}(n)$ of an algorithm A for inputs of length n by $t_A^{\mathrm{av},\mu}(n) = \Sigma_{x \in \Sigma^n} (\mu(x)/\mu(\Sigma^n))\ \mathrm{time}_A(x)$.

Definition 1 ([4]) *A measure μ is malign if it satisfies the following two conditions for any step counting function*: (i) $\mu(\Sigma^n) > 0$ *for any* n; (ii) *for any algorithm A that halts for all inputs there exists a constant c (> 0) such that $t_A^{\mathrm{wo}}(n) \leq c\, t_A^{\mathrm{av},\mu}(n)$ for any n.*

3 A Priori Measures

There are several equivalent ways to define a priori measures. The first is to define them as semicomputable measures that are largest among all semicomputable measures.

Definition 2 *An a priori measure is a semicomputable measure μ such that for any semicomputable measure μ' there exists a constant $c(>0)$ such that $\mu'(x) \leq c\mu(x)$ for any x.*

Existence of such measures is well-known ([3]).

Other ways to define a priori measures use prefix-free algorithms. For two words x, y, y is a *prefix* of x if there exists a word z such that $yz = x$. A set of words L is *prefix-free* if it contain no two different words x, y such that y is a prefix of x. Note that the value $\Sigma_{x \in L} 2^{-|x|}$ is always defined and is at most 1 for any prefix-free set L.

An algorithm A is *prefix-free* if the domain $\mathrm{dom}(\phi_A)$ of the partial function ϕ_A is prefix-free. It is well-known that there exist one prefix-free algorithm A_{U} and one prefix-free set of words L_{U} that satisfy the condition: for any prefix-free algorithm A there exists $h \in L_{\mathrm{U}}$ such that $\phi_{A_{\mathrm{U}}}(hx) = \phi_A(x)$ for any x. We call such A_{U} a *universal prefix-free algorithm* and the word h a *code* of the algorithm A for A_{U}. We use A_{U} and L_{U} to denote some fixed universal prefix-free algorithm and set of codes.

For each word x we define a real number $P(x)$ and a natural number $H(x)$ by $P(x) = \Sigma_y \{2^{-|y|} \mid \phi_{A_{\mathrm{U}}}(y) = x\}$ and $H(x) = \min\{|y| \mid y \in \Sigma^*, \phi_{A_{\mathrm{U}}}(y) = x\}$. (We use the notation $\Sigma_x \{f(x) \mid S(x)\}$ to denote the sum of $f(x)$ for all words x that satisfy the condition $S(x)$.) The value $P(x)$ is the probability that a randomly selected infinite sequence of 0, 1 starts with a word y such that $\phi_{A_{\mathrm{U}}}(y)=x$. The value $H(x)$ is the value that is known as the (prefix-free algorithm based) *Kolmogorov complexity* of x ([3]). The second way to define a priori measures μ is by the condition "μ is semicomputable and $\mu(x) \approx P(x)$" and the third is by the condition "μ is semicomputable and $\mu(x) \approx 2^{-H(x)}$." These three definitions of a priori measures are equivalent. From now on we fix one a priori measure and denote it by $\tilde{\mu}$.

Theorem 1 ([2]) *If μ is an a priori measure then it is malign.*

Proof. The proof in [2] is quite simple. However, here we devide it into two parts to make its logical structure clearer. The proof of the condition $\mu(\Sigma^n) > 0$ is easy and we omit it. Let μ be an a priori measure and let A be an algorithm that halts for all inputs. Suppose that a step counting function is fixed. Let $w(n)$ denote one of the worst inputs of length n.
(1) The proof that there exists a constant $c(> 0)$ such that $\mu(w(n)) \geq c\mu(\Sigma^n)$ for any n. Let A_0 be an algorithm such that

$$\phi_{A_0}(y) = \begin{cases} w(|\phi_{A_U}(y)|) & \phi_{A_U}(y) \text{ is defined,} \\ \text{“undefined”} & \text{otherwise.} \end{cases}$$

This algorithm A_0 is prefix-free. Let h be one of its codes. Then there exist constants c_1, c_2 (> 0) such that we have, for any n, $\mu(w(n)) \geq c_1 P(w(n)) \geq c_1 2^{-|h|} \Sigma_y \{2^{-|y|} \mid \phi_{A_0}(y) = w(n)\}$ $\geq c_1\ 2^{-|h|}\ P(\Sigma^n) \geq c_1\ 2^{-|h|}\ c_2 \mu(\Sigma^n)$. Denoting the constant $c_1 2^{-|h|} c_2$ (> 0) by c, we have $\mu(w(n)) \geq c\mu(\Sigma^n)$ for any n.
(2) The proof that $t_A^{\mathrm{wo}}(n) \leq (1/c) t_A^{\mathrm{av},\mu}(n)$ for any n. For any n, we have $t_A^{\mathrm{av},\mu}(n) = \Sigma_{x \in \Sigma^n}$ $\mathrm{time}_A(x)\ \mu(x)\ /\ \mu(\Sigma^n) \geq \mathrm{time}_A(w(n))\ \mu(w(n))\ /\ \mu(\Sigma^n) \geq c\, t_A^{\mathrm{wo}}(n)$. □

4 Strongly Malign Measures

We defined malignness by Definition 1. However, there are several other ways to define the intuitive notion of malignness. The variations come from two factors.

First factor is what kind of step counting functions are allowed. In [2], this factor was not clearly stated. One extreme way is to fix one step counting function such as the number of steps performed by Turing machines. This is the approach adopted in [4]. Another extreme way is to allow all step counting functions that satisfy the Blum's axioms. Our definition adopts this approach. There are many other ways that are in between these two extreme approaches. At present, we do not know whether the notions of malignness defined by these two extreme approaches are different or not.

The second factor is how to define the sizes of inputs. In Definition 1, the size of an input x meant its length $|x|$. However, in practice we use the term "size of an input" more loosely. For example, if A is an algorithm for computing the product of two Boolean matrices, we will use "an input of size n" to denote a bit sequence of length $2n^2$ that is the representation of two $n \times n$ bit matrices. Generally, the definition of the size of an input depends on the algorithm that uses them as inputs.

There are several ways to formally define malignness. However, it is interesting to note that Li and Vitányi's result that a priori measures are malign seems to hold true for all of these variations of the definition of malignness. Therefore, a priori measures might have one essential feature from which follows the malignness of a priori measure for all these variations.

Analyzing the proof of Li and Vitányi's result (Theorem 1), we see that what they essentially proved is the following property of a priori measures μ.

(A1) There exists a constant $c(> 0)$ such that $\mu(w(n)) \geq c\mu(\Sigma^n)$ for any n.

The malignness of μ immediately follows from (A1). However, Li and Vitányi's proof can be readily modified to prove the following stronger property.

(A2) For any two partial recursive functions f from N to Σ^* and g from Σ^* to N, there exists a constant $c(> 0)$ such that $\mu(f(n)) \geq c\mu(g^{-1}(n))$ for any n such that $f(n)$ is defined.

We think that the porperty (A2) is a feature of a priori measures that deserves investigation in itself and call measures that have this property (A2) "strongly malign."

Definition 3 *A measure μ is strongly malign if it satisfies the following two conditions:*

(1) $\mu(\Sigma^*) > 0$;

(2) *for any algorithms A, B, there exists a constant $c(> 0)$ such that, for any $n \in$ N, if $\phi_A(n)$ is defined then $\mu(\phi_A(n)) \geq c\mu(\phi_B^{-1}(n))$.*

As is expected, a priori measures are strongly malign and strongly malign measures are malign.

Theorem 2 *A priori measures are strongly malign.*

Proof. Let μ be an a priori measure. Suppose that A, B are algorithms. Let A_0 be an algorithm such that $\phi_{A_0}(y)=\phi_A(\phi_B(\phi_{A_U}(y)))$. Then the step (1) of the proof of Theorem 1 becomes a proof of the strong malignness of μ if we replace $w(n)$ with $\phi_A(n)$, $|\phi_{A_U}(n)|$ with $\phi_B(\phi_{A_U}(n))$ and Σ^n with $\phi_B^{-1}(n)$. □

The proof of the following theorem is easy and we omit it.

Theorem 3 *Strongly malign measures are malign.*

Let AP, SSMAL, SMAL, MAL denote the following classes of measures: AP=the class of all a priori measures, SSMAL=the class of all semicomputable strongly malign measures, SMAL=the class of all strongly malign measures, MAL=the class of all malign measures. The above theorems imply AP $\subseteq$ SSMAL $\subseteq$ SMAL $\subseteq$ MAL. We will show that these four classes are all different. To show SSMAL$\neq$SMAL and SMAL$\neq$MAL is straightforward.

Theorem 4 *There exists a strongly malign measure that is not semicomputable.*

Proof (an outline). For each n let x_n be a word of length n such that $H(x_n) \geq n/2$ and let X be the set $\{x_0, x_1, \ldots\}$. We can relativize the notions of "a priori", "strongly malign" and $H(x)$ with respect to X. Let μ be a measure that is a priori relative to X. Then μ is strongly malign relative to X and is hence strongly malign. Let $\tilde{\mu}$ be an a priori measure. Then we can prove that there exist constants c_1, c_2 such that $\tilde{\mu}(x_n) \leq c_1 2^{-n/2}$ and $\mu(x_n) \geq c_2 2^{-H^X(x_n)} \geq c_2 2^{-2\log n} = c_2/n^2$ for any n ($H^X(x)$ denotes the relativized $H(x)$). This shows that μ is not semicomputable. □

Theorem 5 *There exists a malign measure that is not strongly malign.*

Proof (an outline). Let $\tilde{\mu}$ be an a priori measure. We can easily show that the measure μ defined by $\mu(x) = \tilde{\mu}(x)/|x|$ is malign and is not a priori. □

5 Characterizations of a Priori and Strongly Malign Measures

Before proceeding to the proof of AP $\neq$ SSMAL we give some characterizations of the two classes AP and SMAL. These characterizations show an essential difference, if any, of the two classes AP and SSMAL.

Theorem 6 *For a measure μ, the following two conditions are equivalent.*

(1) *μ is a strongly malign measure.*

(2) *μ satisfies the conditions:*

(i) $\mu(\Sigma^*) > 0$;

(ii) *for any algorithm A there exists a constant $c(> 0)$ such that $\mu(x) \geq c\mu(\phi_A^{-1}(x))$ for any x.*

(I) Proof of (1)⇒(2). Suppose that μ is strongly malign. Let A be an algorithm. Let B, C be algorithms such that $\phi_B(n) = n$, $\phi_C(y) = \phi_A(y)$. Then there exists a constant c (> 0) such that $\mu(\phi_B(n)) \geq c\mu(\phi_C^{-1}(n))$ for any $n \in \mathrm{N}$, and hence $\mu(x) = \mu(\phi_B(x)) \geq c\mu(\phi_C^{-1}(x))$ $= c\mu(\phi_A^{-1}(x))$ for any $x \in \Sigma^*$.

(II) Proof of (2)⇒(1). Suppose that a measure μ satisfies (i) – (ii). Let A, B be algorithms. Let C be an algorithm such that $\phi_C(y) = \phi_A(\phi_B(y))$ for any $y \in \Sigma^*$. There exists a constant $c(> 0)$ such that $\mu(x) \geq c\mu(\phi_C^{-1}(x))$ for any $x \in \Sigma^*$. Then we have $\mu(\phi_A(n)) \geq$ $c\mu(\phi_C^{-1}(\phi_A(n))) \geq c\mu(\phi_B^{-1}(n))$ for any $n \in \mathrm{N}$ such that $\phi_A(n)$ is defined. □

Recall that by $\tilde{\mu}$ we denote some fixed a priori measure.

Theorem 7 *For a measure μ, the following two conditions are equivalent.*

(1) *μ is an a priori measure.*

(2) *μ satisfies the conditions:*

(i) μ *is semicomputable*;

(ii) $\mu(\Sigma^*) > 0$;

(iii) *there exists a constant* $c(> 0)$ *such that* $\mu(x) \geq c\tilde{\mu}(A)\mu(\phi_A^{-1}(x))$ *for any algorithm* A *and any* x.

(I) Proof of (1) $\Rightarrow$ (2). Let μ be an a priori measure. It is obvious that (i), (ii) hold true. Let B be a prefix-free algorithm such that $\phi_B(uv) = \phi_w(\phi_{A_U}(v))$, with $w = \phi_{A_U}(u)$, and let h be one of its codes. If u is a description of A (that is, a word u such that $\phi_{A_U}(u) = A$) and v is a description of a word in $\phi_A^{-1}(x)$, then we have $\phi_{A_U}(huv) = \phi_B(uv) = x$. Therefore there exist constants $c_1, c_2(> 0)$ such that

$$\begin{aligned}
\mu(x) &\geq c_1 P(x) = c_1 \Sigma_y \{2^{-|y|} \mid \phi_{A_U}(y) = x\} \\
&\geq c_1 2^{-|h|} (\Sigma_u \{2^{-|u|} \mid \phi_{A_U}(u) = A\})(\Sigma_v \{2^{-|v|} \mid \phi_{A_U}(v) \in \phi_A^{-1}(x)\}) \\
&= c_1 2^{-|h|} P(A) P(\phi_A^{-1}(x)) \geq c_1 2^{-|h|} c_2 \tilde{\mu}(A) \mu(\phi_A^{-1}(x)).
\end{aligned}$$

(II) Proof of (2)$\Rightarrow$(1) (An outline). Suppose that a measure μ satisfies (i) – (iii). For each word x_0 let A_{x_0} be an alrogithm such that $\phi_{A_{x_0}}(x) = x_0$ for any x. The algorithm A_{x_0} is determined by x_0 and hence there is a constant c_1 such that $H(A_{x_0}) \leq H(x_0) + c_1$. Hence there is a constant $c_2(> 0)$ such that $\tilde{\mu}(A_{x_0}) \geq c_2 \tilde{\mu}(x_0)$, and we have $\mu(x_0) \geq c\tilde{\mu}(A_{x_0})\mu(\phi_{A_{x_0}}^{-1}(x_0)) \geq cc_2\tilde{\mu}(x_0)\mu(\Sigma^*)$. Denoting the constant $cc_2\mu(\Sigma^*)$ (> 0) by c_3, we have $\mu(x_0) \geq c_3\tilde{\mu}(x_0)$ for any x_0. Moreover μ is semicomputable. Hence μ is a priori. □

Comparing these two theorems, we see that the essential difference between SSMAL and AP lies in the difference of the behavior of $\mu(x)/\mu(\phi_A^{-1}(x))$ as a function of x. For a measure in SSMAL, this value is bounded from below by a constant c_A (> 0) that may depend on the algorithm A. For a measure in AP, the same is true and moreover the constant c_A can be represented as $c\tilde{\mu}(A)$ with a constant c (> 0) that does not depend on A. Therefore, the problem of whether AP = SSMAL or not is the problem of whether the nonuniform existence of constants for measures in SSMAL can be always changed to uniform existence

or not.

6 Semicomputable Strongly Malign Measures that are not a Priori

We show one method for constructing a family of semicomputable strongly malign measures. This family contains some of a priori measures, and also measures that are not a priori.

Let f be a computable function from $\Sigma^* \times \Sigma$ to Q such that $f(u,a) > 0$, $f(u,0)+f(u,1) \leq 1$, and $f(uv,a) \geq f(v,a)$ for any $u, v \in \Sigma^*$ and $a \in \Sigma$. For this f, let π_f be the function from Σ^* to Q defined by $\pi_f(\lambda) = 1$ and $\pi_f(a_1 a_2 \ldots a_m) = f(\lambda,a_1) f(a_1,a_2) f(a_1 a_2,a_3) \ldots f(a_1 a_2 \ldots a_{m-1}, a_m)$ for $m \geq 1$. It is easy to show that $\pi_f(u_1 \ldots u_m) \geq \pi_f(u_1) \ldots \pi_f(u_m)$ for any words $u_1, \ldots, u_m$ and that $\Sigma_s\{\pi_f(s) \mid s \in L\} \leq 1$ for any prefix-free set L. Finally, let P_f be the measure defined by $P_f(x) = \Sigma_s\{\pi_f(s) \mid \phi_{A_U}(s) = x\}$.

For two rational numbers p, q such that $0 < p$, $0 < q$, $p + q \leq 1$, let $\pi_{q,p}(u)$ and $P_{q,p}(x)$ denote $\pi_f(u)$ and $P_f(x)$ respectively for the function f defined by $f(u,0) = q$, $f(u,1) = p$. For a rational number p such that $0 < p < 1$, let $\pi_p(u)$ and $P_p(x)$ denote $\pi_{1-p,p}(u)$ and $P_{1-p,p}(x)$ respectively. Moreover, for each $a \in \Sigma$ let $|0|_p$ and $|1|_p$ denote $-\log(1-p)$ and $-\log p$ respectively, for each word $u = a_1 \ldots a_m$ let $|u|_p$ denote $|a_1|_p + \ldots + |a_m|_p$ $(= -\log \pi_p(u))$, and for each word x let $H_p(x)$ denote $\min\{|u|_p \mid u \in \Sigma^*, \phi_{A_U}(u) = x\}$. In these notations, the measure $P(x)$ and the Kolmogorov complexity $H(x)$ introduced in Section 3 are $P_{1/2}(x)$ and $H_{1/2}(x)$ respectively.

In this section we show four results: (1) $P_f(x)$ is semicomputable and strongly malign; (2) if $p + q < 1$ then $P_{q,p}(x)$ is not a priori; (3) $P_p(x)$ is a priori; (4) there is a constant c such that $|H(x) - H_p(x)| \leq c$ for any x. The results (1), (2) imply AP $\neq$ SSMAL.

Theorem 8 *The measure $P_f(x)$ is semicomputable and strongly malign.*

Proof. It is obvious that $P_f(x)$ is semicomputable. We use Theorem 6 to prove that $P_f(x)$ is strongly malign. It is obvious that $P_f(\Sigma^*) > 0$. Let A be an algorithm. Let A_0

be a prefix-free algorithm such that $\phi_{A_0}(s) = \phi_A(\phi_{A_U}(s))$ and let h be one of its codes. Then, for any x we have $P_f(x) = \Sigma_r\{\pi_f(r) \mid \phi_{A_U}(r) = x\} \geq \Sigma_s\{\pi_f(hs) \mid \phi_{A_U}(hs) = x\} \geq \pi_f(h)\Sigma_s\{\pi_f(s) \mid \phi_{A_U}(s) \in \phi_A^{-1}(x)\} = \pi_f(h)P_f(\phi_A^{-1}(x))$. Denoting the constant $\pi_f(h)$ (> 0) by c, we have $P_f(x) \geq cP_f(\phi_A^{-1}(x))$. □

For a word $w = a_1a_2 \ldots a_n$ of length n and i such that $0 \leq i \leq n$, let $w|i$ denote the head $a_1a_2 \ldots a_i$ of w of length i.

Theorem 9 *If f satisfies the condition*

$$\lim_{n\to\infty} \max_{w\in\Sigma^n} \prod_{i=0}^{n-1} (f(w|i, 0) + f(w|i, 1)) = 0,$$

then the measure $P_f(x)$ is not a priori.

Proof. Let g be the function from $\Sigma^* \times \Sigma$ to Q defined by $g(u, a) = f(u, a)/(f(u, 0) + f(u, 1))$. Suppose that $P_f(x)$ is a priori. Then there exists a constant $c(> 0)$ such that $P_g(x) < cP_f(x)$ for any x because $P_g(x)$ is semicomputable.

Let n_0 be a value such that $\prod_{i=0}^{n-1}(f(w|i, 0) + f(w|i, 1)) \leq 1/c$ for any $n \geq n_0$ and any word w of length n. For any word $s = a_1a_2 \ldots a_n$ of length $n \geq n_0$ we have $\pi_f(s) \leq (1/c)\pi_g(s)$ because

$$\begin{aligned} \pi_f(s) &= \prod_{i=1}^{n} f(a_1 \ldots a_{i-1}, a_i) = \pi_g(s) \prod_{i=1}^{n} (f(a_1 \ldots a_{i-1}, 0) + f(a_1 \ldots a_{i-1}, 1)) \\ &= \pi_g(s) \prod_{j=0}^{n-1} (f(s|j, 0) + f(s|j, 1)) \leq (1/c)\pi_g(s). \end{aligned}$$

Let x be a word such that there is no s of length less than n_0 such that $\phi_{A_U}(s) = x$. Then we have $P_f(x) = \Sigma_s \{\pi_f(s) \mid \phi_{A_U}(s) = x\} \leq \Sigma_s\{(1/c)\pi_g(s) \mid \phi_{A_U}(s) = x\} = (1/c)P_g(x)$. This contradicts $P_g(x) < cP_f(x)$. □

Corollary 1 *If p, q are rational numbers such that $0 < p$, $0 < q$, $p + q < 1$, then $P_{q,p}(x)$ is a semicomputable strongly malign measure that is not a priori.*

Proof. We have $\lim_{n\to\infty} \max_{w\in\Sigma^n} \prod_{i=0}^{n-1}(f(w|i,0)+f(w|i,1)) = \lim_{n\to\infty} (p+q)^n = 0$. □

Now we show the results (3), (4). In the remainder of this section we fix one rational number p such that $0 < p < 1$ and denote $1-p$ by q. Moreover, we simply write $\pi(u)$ instead of $\pi_p(u)$ $(= \pi_{1-p,p}(u))$.

We define "the interval represented by a word u" as follows. The empty word λ represents the unit interval $(0,1)$. If a word u represents an interval (s,t), then $u0$ represents the interval $(s,\ s+q(t-s))$ and $u1$ represents the interval $(s+q(t-s),\ t)$. It is obvious that the size $t-s$ of the interval (s,t) represented by a word u is $\pi(u)$. Let τ be the function from Σ^* to Q such that the interval represented by u is $(\tau(u),\ \tau(u)+\pi(u))$. This function τ is defined inductively by $\tau(\lambda)=0$, $\tau(u0)=\tau(u)$, $\tau(u1)=\tau(u)+\pi(u)q$. It is easy to show that if $|u|=|u'|$ and u' is greater than u by one as a binary number then $\tau(u)+\pi(u)=\tau(u')$. We call a number r in the unit interval $(0,1)$ *p-rational* if there exists a word u such that $r=\tau(u)$, and *p-irrational* otherwise. We call a subinterval (s,t) of $(0,1)$ a *p-interval* if it is the interval represented by some word u. It is well-known that any subinterval (s,t) of size d $(=t-s)$ of the unit interval $(0,1)$ contains a 1/2-interval (a "binary interval") of size at least $d/4$. This fact is used to prove $2^{-H(x)} \gtrsim P(x)$. We show a similar result for p-intervals.

Lemma 1 *Any subinterval of size d of $(0,1)$ contains a p-interval of size at least $(d/2)\cdot\min\{q,p\}$.*

Proof (an outline). Let (s,t) be a subinterval of size d of $(0,1)$. We may assume that $0<s$ and s,t are p-irrationals. Let u be the shortest word such that $\tau(u)$ is in the interval (s,t). There is a word u' such that $u=u'1$. At least one of $t-\tau(u)$, $\tau(u)-s$ is at least $d/2$. If $t-\tau(u)\geq d/2$ then let h be the value such that $\tau(u)+\pi(u0^h)\geq t$ and $\tau(u)+\pi(u0^{h+1})<t$. Then the p-interval $(\tau(u0^{h+1}),\ \tau(u0^{h+1})+\pi(u0^{h+1}))$ satisfies the condition. If $\tau(u)-s\geq d/2$ then let h be the value such that $\tau(u'01^h)\leq s$, $\tau(u'01^{h+1})>s$. Then the p-interval $(\tau(u'01^{h+1}),\ \tau(u'01^{h+1})+\pi(u'01^{h+1}))$ satisfies the condition. □

Theorem 10 *If p is a rational number such that $0 < p < 1$ then there is a constant c such that $H_p(x) \leq -\log P(x) + c$ for any x.*

Proof (an outline). It is obvious that $P_p(x)$ is semicomputable. We show that there exists a constant $c(> 0)$ such that $P(x) \leq cP_p(x)$ for any x.

Let X be the recursively enumerable set $\{(x, m) \mid x \in \Sigma^*, m \in \mathrm{N}, 2^{-m} < P(x)/2\}$ and let (x_0, m_0), (x_1, m_1), (x_2, m_2), ... be the enumeration of all elements of X by some algorithm. We have $\Sigma_{i=0}^{\infty} 2^{-m_i} \leq 1$ and hence the interval $(2^{-m_0} + \ldots + 2^{-m_{i-1}}, 2^{-m_0} + \ldots + 2^{-m_i})$ is a subinterval of size 2^{-m_i} of the unit interval $(0, 1)$ for each $i \geq 0$. Let z_i be (the leftmost of) the largest p-interval contained in this interval. Let this p-interval z_i be represented by a word u_i. By Lemma 1, the size $\pi_p(u_i)$ of the p-interval z_i is at least $(2^{-m_i}/2)\min\{q, p\}$. Let A be the algorithm such that if there exists i such that $u = u_i$ then $\phi_A(u) = x_i$, and $\phi_A(u)$ is undefined otherwise. The domain of ϕ_A is the prefix-free set $\{u_0, u_1, \ldots\}$ and hence A is prefix-free. Let h be one of its codes.

Let x be a word. We have $2^{-\lceil -\log P(x)\rceil - 2} < P(x)/2$. Hence $(x, \lceil -\log P(x)\rceil + 2)$ is in X. Let (x_i, m_i) be this element. Then we have $\phi_{A_U}(hu_i) = \phi_A(u_i) = x_i = x$, and hence $H_p(x)$ $\leq |hu_i|_p = -\log \pi(hu_i) = -\log \pi(h) - \log \pi(u_i) \leq -\log \pi(h) + m_i + 1 - \log(\min\{q, p\})$ $\leq -\log \pi(h) + 4 - \log(\min\{q, p\}) - \log P(x)$. Denoting the constant $-\log \pi(h) + 4 -$ $\log(\min\{q, p\})$ by c, we have $H_p(x) \leq -\log P(x) + c$. □

The proofs of the following two corollaries are easy if we note that $2^{-H_p(x)}$ is semicomputable and that $P_p(x) \geq 2^{-H_p(x)}$.

Corollary 2 *If p is a rational number such that $0 < p < 1$, then $P_p(x)$ is an a priori measure.*

Corollary 3 *If p is a rational number such that $0 < p < 1$, then there exists a constant c such that $|H(x) - H_p(x)| \leq c$ for any x.*

7 One Condition for Semicomputable Strongly Malign Measures to be a Priori

In this section we show one condition for measures to be a priori under the assumption that they are semicomutable and strongly malign. We expect such a condition to be weaker than the conditions for the case where no assumption on the measures is made.

We call a function F from Σ^* to R *feasible* if it is semicomputable, $F(s) \geq 0$ for any word s, $F(L) \leq 1$ for any prefix-free set L, and there exists a prefix-free algorithm A such that the measure $\mu(x)$ defined by $\mu(x) = \Sigma_s\{F(s) \mid \phi_A(s) = x\}$ is a priori. The function π_p is an example of feasible functions by Corollary 2. Especially the function $F(s) = 2^{-|s|}$ is feasible. For a measure μ, let $\rho_\mu(s)$ denote $\inf_{y\in\Sigma^*} \mu(sy)/\mu(y)$.

Theorem 11 *Let F be a feasible function and μ be a semicomputable strongly malign measure. Then the following two conditions are equivalent.*

(1) *μ is an a priori measure.*

(2) *For any recursively enumerable prefix-free set L there exists a constant $c(>0)$ such that $F(s) \leq c\rho_\mu(s)$ for any s in L.*

Proof of (1)⇒(2) (an outline) Let L be a recursively enumerable prefix-free set. Using an enumeration (s_0, m_0), (s_1, m_1), ...of the recursively enumerable set $X = \{(s, m) \mid s \in L, m \in \mathrm{N},\ 2^{-m} < F(s)/2\}$, we can show that there is a prefix-free algorithm A_0 such that for any $s \in L$ there exists u such that $\phi_{A_0}(u) = s$ and $|u| \leq -\log F(s) + 5$. Its proof is similar to that of Theorem 10, but we use Lemma 1 with $p = 1/2$.

Let A_1 be a prefix-free algorithm such that $\phi_{A_1}(uv) = \phi_{A_0}(u)\phi_{A_U}(v)$ and let h be one of its codes. Then, for any $s \in L$ and any y we have $H(sy) \leq |h| - \log F(s) + 5 + H(y)$, and hence $\mu(sy) \geq c_1 2^{-H(sy)} \geq c_1 2^{-|h|-5} F(s) 2^{-H(y)} \geq c_1 c_2 2^{-|h|-5} F(s)\mu(y)$ for some constants c_1, c_2 (> 0).

Proof of (2)⇒(1). Let A_0 be a prefix-free algorithm such that the measure $\mu'(x)$ defined by $\mu'(x) = \Sigma_s\{F(s) \mid \phi_{A_0}(s) = x\}$ is a priori. The domain $\mathrm{dom}(\phi_{A_0})$ of ϕ_{A_0} is a recursively

enumerable prefix-free set. Let $c_1(>0)$ be a constant such that $F(s) \leq c_1\rho_\mu(s)$ for any $s \in \mathrm{dom}(\phi_{A_0})$. Let A_1 be the algorithm such that $\phi_{A_1}(sy) = \phi_w(y)$, with $w = \phi_{A_0}(s)$. By Theorem 6 there exists a constant $c_2(>0)$ such that $\mu(x) \geq c_2\mu(\phi_{A_1}^{-1}(x))$ for any x.

For any algorithm A and any word x, we have

$$\begin{aligned}
\mu(\phi_A^{-1}(x))\mu'(A) &= \mu(\{y \mid y \in \Sigma^*, \phi_A(y) = x\})\Sigma_s\{F(s) \mid \phi_{A_0}(s) = A\} \\
&\leq \Sigma_{s,y}\{\mu(y)F(s) \mid s \in \mathrm{dom}(\phi_{A_0}), \phi_{A_1}(sy) = x\} \\
&\leq \Sigma_{s,y}\{c_1\mu(sy) \mid s \in \mathrm{dom}(\phi_{A_0}), \phi_{A_1}(sy) = x\} \\
&\leq c_1\mu(\phi_{A_1}^{-1}(x)) \leq c_1(1/c_2)\mu(x).
\end{aligned}$$

This shows that μ is a priori by Theorem 7. □

References

1. M. Blum: A machine-independent theory of the complexity of recursive functions. J. ACM 14, 322-336 (1967)

2. M. Li, P.M.B. Vitányi: A theory of learning simple concepts under simple distributions and average case complexity for the universal distribution. In: Proc. of the 30th FOCS, 1989, pp.34-39

3. M. Li, P.M.B. Vitányi: Kolmogorov complexity and its applications. In: J. van Leeuwen (ed.): Handbook of Theoretical Computer Science, Vol. A, Chap. IV. Elsevier and MIT Press 1990, pp.187-254

4. P.B. Miltersen: The complexity of malign ensembles. In: Proc. of the 6th SICT, 1991, pp.164-171